

Higher Derivative Gravity and Conformal Gravity From Bimetric and Partially Massless Bimetric Theory

S.F. Hassan, Angnis Schmidt-May, Mikael von Strauss

*Department of Physics & The Oskar Klein Centre,
Stockholm University, AlbaNova University Centre, SE-106 91 Stockholm, Sweden*

E-mail: fawad@fysik.su.se, angnis.schmidt-may@fysik.su.se,
mvs@fysik.su.se

ABSTRACT: In this paper we establish the correspondence between ghost-free bimetric theory and a class of higher derivative gravity actions, including conformal gravity and New Massive Gravity. We also characterize the relation between the respective equations of motion and classical solutions. We illustrate that, in this framework, the spin-2 ghost of higher derivative gravity is an artifact of the truncation to a 4-derivative theory. The analysis also gives a relation between the proposed partially massless (PM) bimetric theory and conformal gravity, showing, in particular, the equivalence of their equations of motion at the 4-derivative level. For the PM bimetric theory this provides further evidence for the existence of an extra gauge symmetry and the associated loss of a propagating mode away from de Sitter backgrounds. The new symmetry is an extension of Weyl symmetry which also suggests the PM bimetric theory as a ghost-free completion of conformal gravity.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | A review of the different theories considered | 1 |
| 1.2 | Summary of results | 4 |
| 2 | Higher curvature gravity from bimetric theory | 6 |
| 2.1 | Outline of obtaining higher derivative gravity from bimetric theory | 6 |
| 2.2 | Review of ghost free bimetric gravity | 8 |
| 2.3 | The algebraic solutions for S and f | 9 |
| 2.4 | Higher derivative gravity from bimetric theory | 11 |
| 3 | The ghost issue and relevance to New Massive Gravity (NMG) | 14 |

| | | |
|----------|--|-----------|
| 4 | Conformal gravity from partially massless bimetric theory | 15 |
| 4.1 | The correspondence in $d = 4$ | 15 |
| 4.2 | A step further: Equivalence between CG and PM bimetric theory | 16 |
| 5 | Discussions | 18 |
| A | Higher derivative treatment of free massive spin-0 and spin-2 fields | 19 |
| A.1 | Higher derivative treatment of massive scalars | 20 |
| A.2 | Higher derivative treatment of linearized bimetric theory | 23 |
| B | The general perturbative solution of the $g_{\mu\nu}$ equation for $f_{\mu\nu}$ | 25 |

1 Introduction

In this paper we show the correspondence between the ghost-free bimetric theory [1, 2] and higher derivative gravity, both of which have similar spectra, but only the bimetric case is ghost-free. In particular, this implies a close relation between conformal gravity and a specific bimetric theory that has been proposed [3, 4] as the sought-after nonlinear partially massless (PM) theory. In this section, we start with a brief discussion of various theories considered in this paper, emphasizing the features that are of relevance here. We then present a summary of our results. In section 2, we consider the correspondence for general bimetric parameters and the ghost issue is discussed in section 3. Section 4 is devoted to the PM bimetric theory and conformal gravity. Section 5 contains some discussions. A scalar field example is worked out in appendix A and some calculational details are relegated to appendix B.

1.1 A review of the different theories considered

Below we briefly review the relevant features of higher derivative gravity, conformal gravity, ghost-free bimetric theory, and partially massless theories.

Higher derivative (HD) gravity:

By this we mean theories with more than two derivatives of the metric $g_{\mu\nu}$ that, at the 4-derivative level, have the form,

$$S_{(2)}^{\text{HD}}[g] = m_g^2 \int d^4x \sqrt{g} \left[\Lambda + c_R R(g) - \frac{c_{RR}}{m^2} \left(R^{\mu\nu} R_{\mu\nu} - \frac{1}{3} R^2 \right) \right]. \quad (1.1)$$

This action propagates a massless spin-2 state with two helicities, along with a massive spin-2 state with five helicities, for a total of seven modes [5, 6]. The massive spin-2 state of mass $m_2^2 = m^2 c_R / c_{RR}$ is a ghost, hence it violates unitarity. Changing the coefficient of the R^2 term away from $\frac{1}{3}$ adds a massive scalar to the spectrum. Precisely for the form (1.1) the

scalar becomes infinitely massive and decouples. To (1.1) one may add higher derivative terms suppressed by higher powers of $1/m^2$. The above action can be generalized to d dimensions,

$$S_{(2)}^{\text{HD}}[g] = m_g^{d-2} \int d^d x \sqrt{g} \left[\Lambda + c_R R(g) - \frac{c_{RR}}{m^2} \left(R^{\mu\nu} R_{\mu\nu} - \frac{d}{4(d-1)} R^2 \right) \right]. \quad (1.2)$$

For $d = 3$ and $m_g c_R < 0$, the action (1.2) has been interpreted as a ghost-free theory of massive gravity, dubbed New Massive Gravity [7]. The choice $m_g c_R < 0$ renders the massive spin-2 state healthy. But the massless state, which is now a ghost, does not propagate in $d = 3$, so the spectrum consists of a single massive spin-2 state. This construction is peculiar to $d = 3$ and can not cure the ghost problem for $d \geq 4$ [8].

Conformal gravity:

Dropping the first two terms in (1.1), one obtains the action for conformal gravity, [9],

$$S^{\text{CG}}[g] = -\frac{c_{RR} m_g^2}{m^2} \int d^4 x \sqrt{g} \left[R^{\mu\nu} R_{\mu\nu} - \frac{1}{3} R^2 \right]. \quad (1.3)$$

This is the square of the Weyl tensor (modulo the Euler density) and is invariant under Weyl scalings of the metric. It can also be constructed as the gauge theory of the conformal group [10]. The new Weyl invariance removes one of the seven propagating modes of (1.1), so (1.3) propagates only six modes [11, 12]. In flat backgrounds these appear as a healthy massless spin-2, a ghost-like massless spin-2 and massless spin-1 [13]. In de Sitter (dS) backgrounds, the helicity 1 modes and one set of the helicity 2 modes become massive [14]. Conformal gravity appears in many different contexts and has played a role in several theoretical developments. For a discussion and brief history see [14–17].

Ghost-free bimetric theory:

This is formulated in terms of two metrics, $g_{\mu\nu}$ and $f_{\mu\nu}$, with an interaction potential described in section 2.2. The bimetric action $S[g, f]$ is ghost-free and propagates seven modes [1, 2] (see also [18]). For proportional metrics, $f_{\mu\nu} = c^2 g_{\mu\nu}$, the classical solutions coincide with those of general relativity and, around such backgrounds, the seven modes split into massless and massive spin-2 fields [19]. Due to the presence of the massless field, this is not a theory of massive gravity. Rather, it describes gravity in the presence of a neutral spin-2 field in a completely dynamical setup. The bimetric spectrum is similar to that of HD gravity (1.1), except that neither of the spin-2 fields is a ghost. For some recent work within bimetric theory, see [20–34].

We refer to the ghost-free bimetric models as the Hassan-Rosen (HR) models, as distinct from the fixed background dRGT models [35, 36] to emphasize that they have very distinct dynamics and physics. The dRGT models, that propagate five nonlinear modes [37, 38], are obtainable from the HR ones by dropping both the kinetic term as well as the equation of motion for $f_{\mu\nu}$, which is then set by hand to some fixed background metric, say, $f_{\mu\nu} = \eta_{\mu\nu}$, as in [35, 36]. However, it is important to realize that this procedure cannot be implemented

at the level of the bimetric equations of motion in any reasonable way [39]. Consequently, the HR and dRGT setups not only have different field content, but they also have very distinct dynamics and predictions. In particular, this difference is crucial for understanding that the recent analysis of acausality in massive gravity [40] and the arguments for the absence of partially massless theories in the fixed background setup [41, 42] do not apply to the HR theory. These works rely on the existence of a non-dynamical $f_{\mu\nu}$ and break down in the fully dynamical bimetric setup. Furthermore, since dRGT is not a limit of the HR theory, there is no reason to expect that results specifically derived for the former will automatically have implications for the latter, in the absence of any further evidence. Hence, to reiterate, the analysis of [40, 41] in the dRGT context, does not apply to the HR bimetric theory.

Linear partially massless (PM) theory and beyond:

Partial masslessness was first observed in the Fierz-Pauli theory of a linear massive spin-2 field in a dS background. In $d = 4$ it was found that when the mass and the cosmological constant satisfy the Higuchi bound, $m_{FP}^2 = \frac{2}{3}\Lambda$ [43], the massive spin-2 field has four propagating modes, instead of the usual five [44]. The reduction is due to a new gauge symmetry that emerges at the Higuchi bound. The resulting theory has been called partially massless. The obvious question is if there exists a nonlinear PM theory that goes beyond the Fierz-Pauli setup and that could even give a meaning to partial masslessness away from dS backgrounds.¹

A perturbative approach to finding such a theory has been to construct cubic vertices with PM gauge invariance in the linear FP theory [47, 48]. This has led to the insight that, in a two-derivative theory, cubic PM vertices exist only in 3 and 4 dimensions [46, 49]. In higher derivative theories, such vertices can also be constructed in higher dimensions [46]. However, extending this method beyond cubic order becomes rather tedious.

Alternatively, a footnote in [14] observed that conformal gravity (1.3) around dS backgrounds propagates a massless spin-2 field along with 4 massive modes that could be identified with the PM field. One of the two fields is now a ghost, depending on the overall sign of the action. The possibility of conformal gravity as a nonlinear PM theory was further investigated in [50]. The criterion was to look for a 4-component PM field away from dS or Einstein backgrounds.² Such modes were not found leading to the conclusion that conformal gravity is not a nonlinear PM theory, though it could still prove useful to the problem. From the bimetric point of view described below, the criterion in [50] is too restrictive as it requires that the PM theory around any background looks similar to that around a dS background.

The main obstacle to a systematic construction of nonlinear PM theory is the relation to massive spin-2 fields that generically contain a Boulware-Deser ghost [51]. Hence the search must focus on such ghost-free theories. The first of these is the dRGT model [35, 36], often described as massive gravity, though it also describes a massive spin-2 field $g_{\mu\nu}$ in a flat spacetime $\eta_{\mu\nu}$. The model is ghost-free nonlinearly [37], but it does not admit dS backgrounds

¹The PM phenomenon also arises in higher spin theories [45–47], but here we consider only the spin-2 case.

²Linear PM modes exist in backgrounds specified by $R_{\mu\nu} = \Lambda g_{\mu\nu}$, i.e., in Einstein spacetimes and not just in dS spacetimes. From now on this is assumed to be understood even when not explicitly stated.

[52] and hence does not accommodate even the linear PM theory. The generalization of dRGT models to theories of a massive spin-2 field $g_{\mu\nu}$ in arbitrary curved backgrounds $f_{\mu\nu}$ [53] also exist and are free of the BD ghost [38, 54]. These do admit dS backgrounds for $\bar{g} = \lambda^2 f$ and can easily accommodate the known linear PM theory. However, it can be seen that the fixed non-dynamical background is inconsistent with the linear PM gauge symmetry [3, 4]. A nonlinear PM candidate along these lines had been proposed in [55] but it was subsequently shown that this theory cannot accommodate an extra nonlinear gauge invariance [41, 42].

Partially massless (PM) bimetric theory:

The HR bimetric theory [1] has dS solutions for $f_{\mu\nu} = c^2 g_{\mu\nu}$, where, in general, c^2 is determined in terms of the parameters β_n and $\alpha = m_f/m_g$ of the theory. In such backgrounds, the expressions for the cosmological constant and the FP mass of the spin-2 fluctuations are known [4, 19]. On imposing the Higuchi bound, one can easily recover the linear PM theory and its gauge symmetry. Then, as it turns out, a necessary condition for the consistency of the linear PM symmetry with the dynamical backgrounds is that the theory leaves c^2 undetermined [3, 4]. This uniquely determines a specific bimetric theory as the candidate nonlinear PM theory that exists even away from dS or Einstein backgrounds.

Although the gauge symmetry away from dS backgrounds is not yet known, there are indications that it exists, in which case the theory will propagate six modes (instead of the seven for a generic HR model) even away from dS backgrounds. However only around dS backgrounds, these can be decomposed as four components of the PM field and two helicities of a massless graviton [19]. Otherwise, the six modes will not have such simple decompositions in terms of mass eigenstates. Such a behaviour is not unexpected from a field theoretical perspective due to the nonlinear interactions of fields with the same spin. Hence, our criterion for identifying a nonlinear generalization of the known PM theory is that it has six propagating modes (instead of the generic seven), irrespective of how they decompose.³

One evidence in favour of the bimetric PM proposal is that, in a two derivative setup, it correctly predicts nonlinear PM theories only in $d = 3$ and $d = 4$. However, on allowing for higher derivative terms one obtains candidate PM theories even in higher dimensions [4]. This is in accordance with the outcome of direct construction of PM cubic vertices mentioned above (for some more arguments, see [3, 4]). Further evidence for the existence of an extra symmetry in the PM bimetric theory will be provided below.

1.2 Summary of results

Our results are summarized below:

Correspondence between bimetric theory and HD gravity: Starting with the bimetric action $S[g, f]$ in terms of spin-2 fields $g_{\mu\nu}$ and $f_{\mu\nu}$, we determine $f_{\mu\nu}$ algebraically in terms of $g_{\mu\nu}$ and its curvatures $R_{\mu\nu}(g)$. In general, the solution $f_{\mu\nu}(g)$ is a perturbative

³In the absence of a nonlinearly identifiable 4-component PM multiplet, one may object to labeling these as PM theories. We use the name at least for historical reasons, but more importantly, since these are seemingly the only possible nonlinear generalizations of the linear PM theories.

expansion in powers of $R_{\mu\nu}(g)/m^2$, where, m^2 sets the scale of the Fierz-Pauli mass in the bimetric theory. While this expression is valid for $R(g) < m^2$, in special cases the series terminates and the solution is exact. Using this solution to eliminate $f_{\mu\nu}$ from the bimetric action, we obtain the action for higher derivative gravity as,

$$S^{\text{HD}}[g] = S[g, f(g)]. \quad (1.4)$$

At the 4-derivative level, this gives $S_{(2)}^{\text{HD}}$ in (1.1). We perform this calculation in d dimensions and hence for $d = 3$ we can obtain a higher derivative extension of the New Massive Gravity setup of [7] as dictated by the bimetric theory.⁴

We also obtain a correspondence between the equations of motion in HD gravity and bimetric theory. These differ by an extra operator $(\delta f/\delta g)$. If $\chi(x)$ solves the differential equation $(\delta f/\delta g)\chi = 0$, then the solutions of the HD gravity equations have the form $g^{\text{HD}} = g^{\text{HD}}(x, \chi(x))$. The classical solutions in bimetric theory are related to these by,

$$g^{\text{BM}}(x) = g^{\text{HD}}(x, \chi = 0). \quad (1.5)$$

Thus, on setting $\chi = 0$, say, through boundary conditions, the classical solutions in the two theories will coincide.

Truncation and ghost in HD gravity: The HD gravity action truncated to quadratic curvature terms, $S_{(2)}^{\text{HD}}$ (1.1), has seven propagating modes that, around appropriate backgrounds, decompose into a massless and a massive spin-2 field. This is similar to the bimetric spectrum except for two important differences. The massive spin-2 field in $S_{(2)}^{\text{HD}}$ is a ghost and its mass differs from the value in the associated bimetric theory. We illustrate that these discrepancies are artifacts of truncating HD gravity to a 4-derivative theory. Resolving the ghost issue also requires using appropriate source couplings, as dictated by the associated bimetric theory.

To make this explicit, in the appendix we consider two examples, the linearized bimetric theory, and a very similar but much simpler theory of two scalar fields. In both cases one of the fields can be eliminated to explicitly obtain a higher derivative action for the remaining field, analogous to the untruncated S^{HD} . This action contains appropriate modified source couplings inherited from the bi-field theory. To check for ghosts, we compute the vacuum persistence amplitude which turns out to be exactly the same as the one computed in the associated bi-field theory. Hence the untruncated HD theory is ghost free and has the right mass poles. However, one can see that if this theory is truncated to 4-derivative terms, the massive field becomes a ghost and its mass shifts away from the correct value.

Conformal gravity and the PM bimetric theory: For the specific bimetric theory that has been identified as the candidate nonlinear PM theory [3, 4], the corresponding HD theory at the 4-derivative level is conformal gravity (1.3). What is more, in this case the correspondence is closer to an equivalence: while in general, the truncated equations in HD

⁴For another attempt to obtain New Massive Gravity from bimetric theory, involving certain scaling limits of the parameters, see [56]. The outcome there is not directly comparable to our results.

gravity and bimetric theory are not the same, in the PM case, the lowest order bimetric equation is identical to the conformal gravity equation, expressed as the vanishing of the Bach tensor,

$$B_{\mu\nu}(g) = 0. \quad (1.6)$$

For the PM bimetric theory the implication is that, in the limit $R(g) \ll m^2$, it propagates six (rather than seven) modes even away from dS backgrounds. The new gauge symmetry that exists in the limit $R(g) \ll m^2$, not only around dS backgrounds, but around any Bach flat background, corresponds to Weyl transformations of $g_{\mu\nu}$. This is a further indication that the full PM bimetric theory indeed has an extra gauge symmetry. Since the Bach equation is now derived from the bimetric action (instead of the CG action), none of the six modes it propagates is a ghost.

Conversely, the above equivalence implies that the PM bimetric theory is a genuine extension of conformal gravity that resolves its ghost problem. In the process the Weyl symmetry of CG is replaced by the PM symmetry of bimetric theory.

From the bimetric perspective, it is obvious that only around proportional dS backgrounds the six modes split into a massless graviton and a 4 component PM field. Away from dS backgrounds the six modes do not admit such a split [19] and, in particular, it is not possible to freeze gravity and retain only the PM field, as attempted in [50].

2 Higher curvature gravity from bimetric theory

Starting with the bimetric theory, we will now obtain the higher derivative gravity actions as an expansion in the inverse of the mass scale. We will explicitly compute quadratic curvature terms after first outlining the procedure and a brief review of the ghost-free bimetric theory.

2.1 Outline of obtaining higher derivative gravity from bimetric theory

The bimetric action $S[g, f]$ involves the two spin-2 fields $g_{\mu\nu}$ and $f_{\mu\nu}$ with a non-derivative interaction potential to be specified later. The corresponding equations of motion,

$$\left. \frac{\delta S}{\delta g^{\mu\nu}} \right|_f = 0, \quad \left. \frac{\delta S}{\delta f^{\mu\nu}} \right|_g = 0, \quad (2.1)$$

are coupled differential equations and include the sources $T_{\mu\nu}^g$ and $T_{\mu\nu}^f$. Eliminating one of the metrics, say $f_{\mu\nu}$, between the two equations leads to a higher derivative equation for $g_{\mu\nu}$. Of course this is just a step in the direction of solving (2.1) and the resulting higher derivative equation is completely within the framework of bimetric theory. But, as will be exhibited for a scalar field in the appendix, it is not straightforward to derive this higher derivative equation directly from a local ghost-free action for $g_{\mu\nu}$ alone. So this manipulation will not lead to standard higher curvature gravity actions.

On the other hand, note that the $g_{\mu\nu}$ equations do not contain derivatives of $f_{\mu\nu}$. Hence they can be solved algebraically to determine $f_{\mu\nu}$ in terms of $g_{\mu\nu}$ and its curvatures, at least

perturbatively. This will yield $f = f(g, T^g)$. Using this to eliminate $f_{\mu\nu}$ in the bimetric action $S[g, f]$ then leads to a higher derivative action for $g_{\mu\nu}$ which precisely coincides with the standard class of higher curvature gravity actions. Before getting into the details, let us clarify the relation between the equations of motion in HD gravity and in bimetric theory.

Suppose one obtains $f = f(g)$ by solving either of the equations in (2.1). This can be used to convert the bimetric action into a higher derivative action $S'[g] = S[g, f(g)]$ for g . Varying S' with respect to g gives,

$$\frac{\delta S'}{\delta g(x)} = \frac{\delta S}{\delta g(x)} \Big|_{f(g)} + \int d^d y \frac{\delta f(y)}{\delta g(x)} \frac{\delta S}{\delta f(y)} \Big|_g = 0. \quad (2.2)$$

The first term corresponds to the usual bimetric g -equation of motion (2.2) while the second term corresponds to the bimetric f equation of motion, multiplied by a Jacobian. Two cases arise depending on how $f_{\mu\nu}$ is determined.

If $f_{\mu\nu}$ is a solution to the f -equation $\frac{\delta S}{\delta f}|_g = 0$, say as, $f = f(g, T^f)$, then the second term vanishes and (2.2) implies exactly the bimetric g -equation. However, this requires solving a differential equation for $f_{\mu\nu}$ (which is not easy to solve in full generality) and the solution depends on the boundary conditions. Hence in this case, the action S' is bi-local and the resulting theory is completely equivalent to the bimetric theory. This approach is not the focus of our attention in the present paper.

Now suppose that we instead use the bimetric g -equation, $\frac{\delta S}{\delta g}|_f = 0$, to algebraically determine $f = f(g, T^g)$. Since the solution is algebraic, $S'[g]$ is a local higher derivative action for g and coincides with $S^{\text{HD}}[g]$ (1.4). Later we will see that, when expanded in powers of curvatures, it reproduces the standard higher derivative gravity action (1.2) at the lowest orders. From (2.2), the equation of motion becomes,

$$\frac{\delta S^{\text{HD}}}{\delta g(x)} = \int d^d y \frac{\delta f(y)}{\delta g(x)} \frac{\delta S}{\delta f(y)} \Big|_g = 0. \quad (2.3)$$

The Jacobian factor is in fact an operator, $\frac{\delta f(y)}{\delta g(x)} = \delta(x - y) \hat{\mathcal{O}}_y$, so the higher derivative g -equation of motion is,

$$\hat{\mathcal{O}} \left(\frac{\delta S[g, f]}{\delta f} \Big|_g \right) = 0, \quad (2.4)$$

where, $f = f(g, T^g)$. This differs from the corresponding bimetric equation by the presence of the extra operator $\hat{\mathcal{O}}$. For functions χ such that $\hat{\mathcal{O}}\chi = 0$, we get,

$$\frac{\delta S[g, f]}{\delta f} \Big|_g = \chi(x), \quad (2.5)$$

Thus, from the bimetric point of view, $\chi(x)$ appears as an unusual source term specified through boundary conditions. Let us denote a solution of this equation by $g^{\text{HD}}(x, \chi(x))$. Thus only for boundary conditions that give $\chi = 0$ the solutions of the higher curvature action $S^{\text{HD}}[g]$ coincide with the solutions of bimetric theory, $g^{\text{BM}}(x) = g^{\text{HD}}(x, \chi = 0)$.

The difference between the bimetric and HD gravity equations, contained in $\hat{\mathcal{O}}$, is easy to isolate formally. However, in the truncated theory, $\hat{\mathcal{O}}$ mixes with $\delta S/\delta f$ and, generically, the truncated equations are not equivalent (an exception being conformal gravity). In particular, the higher derivative theory, when truncated to 4-derivative terms, will contain spin-2 ghosts. But as the scalar field example in the appendix illustrates, these are artifacts of the truncation and of using a naive source coupling instead of the one dictated by the parent bimetric theory. Since the truncation scale is given by the mass scale, the truncated theory is meaningful only at energies below the mass scale. Conversely, the bimetric theory seems to provide a ghost-free completion of quadratic curvature gravity.

2.2 Review of ghost free bimetric gravity

Let us briefly review the relevant equations of HR bimetric theory [1]. More details can be found in [19, 53] and in the appendix of [4]. The bimetric action is,

$$S[g, f] = m_g^{d-2} \int d^d x \left(\sqrt{g} R(g) + \alpha^{d-2} \sqrt{f} R(f) - 2m^2 \sum_{n=0}^d \beta_n e_n(S) \right), \quad (2.6)$$

where S stands for the square-root matrix $S^\mu{}_\nu = (\sqrt{g^{-1}f})^\mu{}_\nu$ and $e_n(S)$ are the elementary symmetric polynomials of the eigenvalues of S . Starting with $e_0(S) = 1$, they are constructed iteratively through,

$$e_n(S) = -\frac{1}{n} \sum_{k=1}^n (-1)^k \text{Tr}(S^k) e_{n-k}(S). \quad (2.7)$$

In d dimensions, $e_d(S) = \det S$ and $e_n(S) = 0$ for $n > d$. We will explicitly need,

$$e_1(S) = \text{Tr} S, \quad e_2(S) = \frac{1}{2} [(\text{Tr} S)^2 - (\text{Tr} S^2)]. \quad (2.8)$$

In (2.6), $\alpha = m_f/m_g$ is the ratio of the Planck masses and β_n are $d+1$ dimensionless free parameters. m^2 , which is degenerate with the overall scale of the β_n , sets the mass scale of the massive mode. The action retains its form under the interchange of g and f due to the properties of the $e_n(S)$ [1]. The equations of motion for $g_{\mu\nu}$ and $f_{\mu\nu}$ are,

$$R^\mu{}_\nu(g) - \frac{1}{2} \delta^\mu_\nu R(g) + m^2 \sum_{n=0}^{d-1} (-1)^n \beta_n \mathbb{Y}_{(n)\nu}^\mu(S) = m_g^{2-d} T^{g\mu}{}_\nu, \quad (2.9)$$

$$R^\mu{}_\nu(f) - \frac{1}{2} \delta^\mu_\nu R(f) + \frac{m^2}{\alpha^{d-2}} \sum_{n=1}^d (-1)^n \beta_n \mathbb{Y}_{(n)\nu}^\mu(S^{-1}) = m_f^{2-d} T^{f\mu}{}_\nu, \quad (2.10)$$

where the matrices $\mathbb{Y}_{(n)}$ are given by (see, for example, [4]),

$$\mathbb{Y}_{(n)}(S) = \sum_{k=0}^n (-1)^k e_k(S) S^{n-k}. \quad (2.11)$$

Around proportional backgrounds, $f_{\mu\nu} = c^2 g_{\mu\nu}$, the degrees of freedom propagated by the action (2.6) split into a massless and a massive spin-2 field. In the present conventions, the Fierz-Pauli mass is given by [19],

$$m_{\text{FP}}^2 = m^2 \left(\frac{1 + (\alpha c)^{d-2}}{(\alpha c)^{d-2}} \right) \sum_{n=1}^{d-1} \binom{d-2}{n-1} c^n \beta_n. \quad (2.12)$$

In general, c is determined in terms of β_n and α through the equality of the cosmological constants, $\Lambda_g = \Lambda_f$. Furthermore, if $\Lambda_g = 0$, then (2.12) gives the mass around flat spacetime.

2.3 The algebraic solutions for S and f

The g -equation (2.9) is algebraic in S^μ_ν and hence in $f_{\mu\nu}$, so it can be used to determine $f_{\mu\nu}$ in terms of $g_{\mu\nu}$, its curvatures and $T^g_{\mu\nu}$. First we consider the minimal β_1 model (with all β_n for $n \geq 2$ vanishing) where the solution can be found exactly and then find a perturbative solution in powers of the curvatures for general β_n .

Exact solution in the β_1 model:

In the β_1 model (where $\beta_2 = \dots = \beta_{d-1} = 0$), the $g_{\mu\nu}$ equation of motion becomes,

$$G^\mu_\nu(g) + m^2 [\beta_0 \delta^\mu_\nu - \beta_1 (S^\mu_\nu - \delta^\mu_\nu \text{Tr} S)] = m_g^{2-d} T^g{}^\mu_\nu, \quad (2.13)$$

where $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$ is the Einstein tensor. The trace of this equation determines $\text{Tr} S$. Then, for $T^g{}^\mu_\nu = 0$ one finds the solution,

$$S^\mu_\nu = -\frac{1}{d-1} \frac{\beta_0}{\beta_1} \delta^\mu_\nu + \frac{1}{m^2 \beta_1} P^\mu_\nu, \quad (2.14)$$

Here, $P^\mu_\nu = g^{\mu\lambda} P_{\lambda\nu}$ is the Schouten tensor of $g_{\mu\nu}$ in d dimensions,⁵

$$P_{\mu\nu} = R_{\mu\nu} - \frac{1}{2(d-1)} g_{\mu\nu} R. \quad (2.15)$$

From S one can easily find $f_{\mu\nu} = g_{\mu\lambda} (S^2)^\lambda_\nu$. The solution has the form of a power series in P/m^2 that terminates at the first order.

The solution for S in the presence of matter sources is easily obtained from (2.14) by the replacement $G_{\mu\nu} \rightarrow G_{\mu\nu} - m_g^{2-d} T^g_{\mu\nu}$, or equivalently by,

$$P^\mu_\nu \rightarrow P^\mu_\nu - m_g^{2-d} \left(T^g{}^\mu_\nu - \frac{1}{d-1} T^g{}^\lambda_\lambda \delta^\mu_\nu \right). \quad (2.16)$$

⁵The conventionally normalized Schouten tensor is $\frac{1}{d-2} P_{\mu\nu}$, but the non-standard normalization of (2.15) is more convenient for our purposes.

Perturbative solution for general β_n :

For general β_n in (2.9), finding an exact solution for $S = \sqrt{g^{-1}f}$ is not feasible. But one can always find a perturbative solution as an expansion in m^{-2} . In the absence of matter sources this corresponds to a derivative expansion in powers of $R_{\mu\nu}(g)/m^2$, or equivalently, in $P_{\mu\nu}(g)/m^2$. In order to find the perturbative solution, we make the general ansatz,⁶

$$S^\mu_\nu = a\delta^\mu_\nu + \frac{1}{m^2} (b_1 P^\mu_\nu + b_2 \text{Tr} P \delta^\mu_\nu) + \frac{1}{m^4} (c_1 P^\mu_\nu{}^2 + c_2 P^\mu_\nu \text{Tr} P + c_3 \text{Tr}(P^2) \delta^\mu_\nu + c_4 (\text{Tr} P)^2 \delta^\mu_\nu) + \mathcal{O}(m^{-6}). \quad (2.17)$$

The coefficients a, b_i and c_i are determined by plugging this into (2.9) and matching the coefficients of different powers of P^μ_ν . The details are given in the appendix. The result can be presented simply in terms of the sums,

$$s_k \equiv \sum_{n=k}^{d-1} \binom{d-1-k}{n-k} \beta_n a^n, \quad k = 0, 1, 2, 3, \dots \quad (2.18)$$

The coefficient a is determined in terms of β_n as a solution of the polynomial equation,

$$s_0 \equiv \sum_{n=0}^{d-1} \binom{d-1}{n} \beta_n a^n = 0. \quad (2.19)$$

Non-real solutions are not excluded as long as the parameters in the final action are real. For the b_i one obtains,

$$b_2 = 0, \quad b_1 = \frac{a}{s_1}. \quad (2.20)$$

where s_1 is defined in (2.18). Finally, the c_i are given by

$$c_1 = -c_2 = \frac{as_2}{s_1^3}, \quad c_3 = -c_4 = -\frac{as_2}{2(d-1)s_1^3}. \quad (2.21)$$

These coefficients exist only if $s_1 \neq 0$ (the case $s_1 = 0$ require separate treatment and implies Ricci flat $g_{\mu\nu}$, c.f. (B.5)) and in terms of them the solution for S^μ_ν to $\mathcal{O}(m^{-4})$ reads,

$$S^\mu_\nu = a\delta^\mu_\nu + \frac{a}{s_1 m^2} P^\mu_\nu + \frac{as_2}{s_1^3 m^4} \left[(P^\mu_\nu{}^2 - P P^\mu_\nu) + \frac{1}{d-1} e_2(P) \delta^\mu_\nu \right] + \mathcal{O}(m^{-6}). \quad (2.22)$$

Note that the higher order curvature terms are suppressed by the mass scale,

$$m^2 s_1 = m^2 \sum_{n=1}^{d-1} \binom{d-2}{n-1} \beta_n a^n. \quad (2.23)$$

⁶This ansatz captures the generic class of solutions where a diagonal $R^\mu_\nu(g)$ implies a diagonal S^μ_ν . One may also find special bimetric solutions where a diagonal $R^\mu_\nu(g)$ corresponds to a non-diagonal S^μ_ν . To see this, note that (2.9) has the generic form $R^\mu_\nu = \sum_r q_r (S^r)^\mu_\nu$, where the scalars q_r are functions of the β_n and of $e_n(S)$. It is possible to construct non-diagonal S such that $\sum_r q_r S^r$ is diagonal. The off-diagonal structure depends on the choice of the β_n , for example, this is not possible in the β_1 model. For solutions of this type see, for example, [57].

This is proportional to the Fierz-Pauli mass (2.12) of the massive spin-2 fluctuation on proportional backgrounds $f_{\mu\nu} = a^2 g_{\mu\nu}$ and with vanishing cosmological constant.⁷ Thus the perturbative expansion becomes more accurate for large Fierz-Pauli mass although the procedure does not really correspond to integrating out the massive mode of the bimetric theory. Rather, we are eliminating $f_{\mu\nu}$ which is not a mass eigenstate.

This procedure can be continued to compute the solution for S^μ_ν to any order in m^{-2} . Here we are mostly interested in the quadratic curvature terms in the higher-derivative action for $g_{\mu\nu}$, and hence it suffices to determine the solution for S^μ_ν to second order. As in the β_1 model, matter sources can be taken into account through the replacement (2.16) in the final solution. It is easy to verify that for $\beta_2 = \dots = \beta_{d-1} = 0$, one reproduces the exact result for the β_1 model obtained above as in this case, $s_n = 0$ for $n \geq 2$.

2.4 Higher derivative gravity from bimetric theory

In this section we eliminate $f_{\mu\nu}$ in the bimetric action using (2.22) to obtain a higher curvature theory for $g_{\mu\nu}$. We explicitly retain only quadratic curvature terms and neglect the source terms which can always be reinstated by the replacement (2.16).

Let us begin by eliminating S from the bimetric interaction potential. For this purpose, it is convenient to first express the potential in terms of,

$$M^\mu_\nu \equiv \frac{1}{a} S^\mu_\nu - \delta^\mu_\nu. \quad (2.24)$$

Then, using the properties of the $e_n(S)$ one has,

$$2m^2 \sqrt{g} \sum_{n=0}^d \beta_n e_n(S) = 2m^2 \sqrt{g} \sum_{n=0}^d \alpha_n e_n(M), \quad (2.25)$$

where we have defined

$$\alpha_n = \sum_{k=n}^d \binom{d-n}{k-n} a^k \beta_k. \quad (2.26)$$

Since M directly starts at order m^{-2} , only $e_n(M)$ with $n \leq 2$ contribute to order m^{-2} in the potential. Explicitly, $e_0(M) = 1$ while $e_1(M)$ and $e_2(M)$ are given by,

$$\begin{aligned} e_1(M) &= \frac{1}{s_1 m^2} e_1(P) - \frac{s_2}{s_1^3 m^4} \frac{d-2}{d-1} e_2(P) + \mathcal{O}(m^{-6}), \\ e_2(M) &= \frac{1}{s_1^2 m^4} e_2(P) + \mathcal{O}(m^{-6}). \end{aligned} \quad (2.27)$$

⁷For backgrounds of the type $f_{\mu\nu} = a^2 g_{\mu\nu}$, the cosmological constant Λ_g in the g -equation is given by s_0 . In the present setup, a must be chosen such that $s_0 = 0$.

Hence the potential is given by,

$$2m^2\sqrt{g}\sum_{n=0}^d\beta_n e_n(S) = 2m^2\sqrt{g}\left(\alpha_0 + \frac{\alpha_1}{s_1 m^2}e_1(P) + \frac{1}{s_1^3 m^4}\left(s_1\alpha_2 - s_2\alpha_1\frac{d-2}{d-1}\right)e_2(P)\right) + \mathcal{O}(m^{-4}). \quad (2.28)$$

Using (2.15), it is straightforward to express $e_n(P)$ in terms of curvatures of $g_{\mu\nu}$,

$$e_1(P) = \frac{d-2}{2(d-1)}R, \quad e_2(P) = \frac{1}{2}\left(\frac{d}{4(d-1)}R^2 - R^{\mu\nu}R_{\mu\nu}\right). \quad (2.29)$$

Then the bimetric potential expressed in terms of curvatures becomes

$$2m^2\sqrt{g}\sum_{n=0}^d\beta_n e_n(S) = 2m^2\sqrt{g}\left[\alpha_0 + \alpha_1\frac{d-2}{d-1}\frac{1}{2s_1 m^2}R + \frac{s_1\alpha_2 - s_2\alpha_1\frac{d-2}{d-1}}{2s_1^3 m^4}\left(\frac{d}{4(d-1)}R^2 - R^{\mu\nu}R_{\mu\nu}\right)\right] + \mathcal{O}(m^{-4}). \quad (2.30)$$

Let us now consider the kinetic term $\sqrt{f}R(f)$. In contrast with the bimetric potential, which produces polynomials only in curvatures of $g_{\mu\nu}$, the kinetic term also produces terms with covariant derivatives acting on curvatures. At quartic order in derivatives, the Bianchi identity can be used to show that the only possible such term is $\nabla^2 R$. However, it will turn out in the following that the coefficient of this term is zero for all β_n . Terms involving derivatives of curvatures only start to appear at cubic order.

The solution for $f_{\mu\nu}$ is obtained from (2.22) using $f_{\mu\nu} = g_{\mu\rho}(S^2)^\rho{}_\nu$ and reads, to first order in curvatures,

$$f_{\mu\nu} = a^2 g_{\mu\nu} + \frac{2a^2}{s_1 m^2}P_{\mu\nu} + \mathcal{O}(m^{-4}). \quad (2.31)$$

Note that in $\sqrt{f}R(f)$ only terms up to this order contribute to the quadratic curvature terms. Expanded to first order in powers of m^{-2} , the inverse of $f_{\mu\nu}$ is then given by

$$(f^{-1})^{\mu\nu} = a^{-2}g^{\mu\nu} - \frac{2}{a^2 s_1 m^2}P^{\mu\nu} + \mathcal{O}(m^{-4}). \quad (2.32)$$

Using the curvature relations (see e.g. [58]),

$$\begin{aligned} R_{\mu\nu}(f) &= R_{\mu\nu}(g) + 2\nabla_{[\mu}C_{\alpha]\nu}{}^\alpha - 2C_{\nu[\mu}{}^\beta C_{\alpha]\beta}{}^\alpha, \\ C_{\mu\nu}{}^\alpha &= \frac{1}{2}(f^{-1})^{\alpha\rho}(\nabla_\mu f_{\nu\rho} + \nabla_\nu f_{\mu\rho} - \nabla_\rho f_{\mu\nu}), \end{aligned} \quad (2.33)$$

we can express the curvatures of $f_{\mu\nu}$ in terms of curvatures and connections of $g_{\mu\nu}$. For the $f_{\mu\nu}$ obtained above this gives,

$$R_{\mu\nu}(f) = R_{\mu\nu}(g) + \frac{2}{s_1 m^2}(\nabla_\mu \nabla_\nu P^\alpha{}_\alpha - \nabla^\rho \nabla_\nu P_{\rho\mu} - \nabla^\rho \nabla_\mu P_{\rho\nu} + \nabla^2 P_{\mu\nu}) + \mathcal{O}(m^{-4}). \quad (2.34)$$

The $C_{\mu\nu}{}^\alpha$ contribute the terms linear in $P_{\mu\nu}$ and in $R(f) = f^{\mu\nu} R_{\mu\nu}(f)$ these drop out due to the usual Bianchi identity,

$$\begin{aligned} R(f) &= a^{-2} R(g) - \frac{2}{a^2 s_1 m^2} P^{\mu\nu} R_{\mu\nu}(g) + \frac{4}{a^2 s_1 m^2} (\nabla^2 P^\alpha{}_\alpha - \nabla^\mu \nabla^\nu P_{\mu\nu}) + \mathcal{O}(m^{-4}) \\ &= a^{-2} R(g) - \frac{2}{a^2 s_1 m^2} \left(R^{\mu\nu} R_{\mu\nu} - \frac{1}{2d-2} R^2 \right) + \mathcal{O}(m^{-4}), \end{aligned} \quad (2.35)$$

where we have used the identity $\nabla^2 P^\alpha{}_\alpha - \nabla^\mu \nabla^\nu P_{\mu\nu} = 0$ which follows from the Bianchi identity. As promised, the terms involving derivatives acting on curvatures have dropped out. Also, using $\det(1 + A) = \sum_n e_n(A)$ and retaining the first two terms, we have,

$$\sqrt{f} = a^d \sqrt{g} \sum_{n=0}^d e_n \left(\frac{P}{s_1 m^2} \right) = a^d \sqrt{g} \left(1 + \frac{1}{2s_1 m^2} \frac{d-2}{d-1} R(g) \right) + \mathcal{O}(m^{-4}). \quad (2.36)$$

Then the kinetic terms become

$$\begin{aligned} &\sqrt{g} R(g) + \alpha^{d-2} \sqrt{f} R(f) \\ &= \sqrt{g} \left[\left(1 + (\alpha a)^{d-2} \right) R(g) - \frac{2(\alpha a)^{d-2}}{s_1 m^2} \left(R^{\mu\nu} R_{\mu\nu} - \frac{d}{4(d-1)} R^2 \right) \right] + \mathcal{O}\left(\frac{1}{m^4}\right). \end{aligned} \quad (2.37)$$

Note that this has the same structure as the potential terms. Finally, combining (2.30) and (2.37) we obtain the higher derivative Lagrangian for $g_{\mu\nu}$ to quadratic order in curvatures,

$$\mathcal{L}_{(2)}^{\text{HD}}(g) = m_g^{d-2} \sqrt{g} \left[\Lambda + c_R R(g) - \frac{c_{RR}}{m^2} \left(R^{\mu\nu} R_{\mu\nu} - \frac{d}{4(d-1)} R^2 \right) \right] + \mathcal{O}(m^{-4}). \quad (2.38)$$

This coincides with (1.1) and the parameters are given by,

$$\begin{aligned} \Lambda &= -2m^2 \alpha_0, \\ c_R &= 1 + (\alpha a)^{d-2} - \frac{\alpha_1}{s_1} \frac{d-2}{d-1}, \\ c_{RR} &= \frac{1}{s_1^2} \left(2(\alpha a)^{d-2} s_1 - \alpha_2 + \frac{s_2}{s_1} \frac{d-2}{d-1} \alpha_1 \right). \end{aligned} \quad (2.39)$$

The theory can also be written in a more compact form,

$$\mathcal{L}_{(2)}^{\text{HD}}(g) = m_g^{d-2} \sqrt{g} \left[\Lambda + c'_R e_1(P) + c'_{RR} e_2(P) \right] + \mathcal{O}(m^{-4}), \quad (2.40)$$

where, $c'_R = 2(d-1)c_R/d-2$ and $c'_{RR} = 2c_{RR}/m^2$.

An identity that follows from (2.18) and (2.26) is,

$$s_n = \alpha_n - \alpha_{n+1}. \quad (2.41)$$

Since $s_0 = 0$, this implies $\alpha_0 = \alpha_1$ and also gives an expression for the coefficients in terms of the α_n , or using (2.26) in terms of the β_n . Note that $\mathcal{L}_{(2)}^{\text{HD}}$ depends on only 3 combinations of the $d + 1$ parameters β_n . To see all the free parameters of the bimetric theory one needs to consider higher order curvature terms in \mathcal{L}^{HD} . The truncation is accurate only in the low curvature regime $R(g) \ll m^2$. It is worth pointing out however, that although the HD Lagrangian is in general an infinite expansion in curvatures, the number of parameters determining this expansion is just the $d + 1$ parameters β_n together with $\alpha = m_f/m_g$.

3 The ghost issue and relevance to New Massive Gravity (NMG)

The higher derivative gravity Lagrangian (2.38) propagates 7 modes in $d = 4$, comprising 2 modes of a healthy massless graviton and 5 modes of a massive spin-2 field which is a ghost. This is in contrast to the original bimetric theory in which all 7 modes are healthy. The masses of the spin-2 fields are also not equal in the two theories. Of course in the HD gravity case, one is working with a truncated action where the mass of the problematic excitation is related to the truncation scale. So it is expected that the ghost problem is an artifact of the truncation.

To illustrate this, in the appendix we consider two examples, the linearized version of bimetric theory, and a very similar but much simpler theory of two scalar fields. In the absence of interactions, these theories can be manipulated explicitly. In both the spin-0 and spin-2 examples, one of the fields can be eliminated to explicitly obtain a closed form higher derivative action for the remaining field, including appropriate source couplings inherited from the original bi-field theory. The action contains up to 6-derivative terms. To check for ghosts, we compute the vacuum persistence amplitude in the higher derivative theory. This turns out to be exactly the same as the one computed in the associated bi-field theory. The equivalence implies that the untruncated HD theory is ghost free and has the right mass poles. However, if this theory is now truncated to 4-derivative terms, the massive field becomes a ghost and its mass shifts away from the correct value, thus illustrating the point.

In the presence of bimetric interactions, the corresponding HD gravity obtained in a perturbative manner will not have a finite number of terms and it is not feasible to write such a theory in a closed form. Considering this, $S_{(2)}^{\text{HD}}$ (1.1) is usable only as an effective theory at energy scales below the mass pole and the R^2 terms must be treated only perturbatively. The present analysis, however, implies that $S_{(2)}^{\text{HD}}$ has a minimal ghost-free completion in the form of an HR bimetric theory. Even simpler, since at the 4-derivative level the HD action has only three independent parameters, its structure can be reproduced by the β_1 model alone. This will also avoid the issue raised in footnote 6. The equations of motion derived from the full S^{HD} differ from the corresponding bimetric equation (obtained on eliminating $f_{\mu\nu}$ between (2.9) and (2.10)) by the extra operator $\delta f/\delta g$, as discussed earlier. The truncation procedure mixes this operator with the HD equation of motion and the similarity with the bimetric equation is no longer obvious.

The correspondence between the HR bimetric theory and the action S^{HD} may also be regarded as a reconfirmation that the bimetric theory propagates only seven modes and no Boulware-Deser ghost [51].

One can also consider the implications of this correspondence to New Massive Gravity (NMG). There has been some work to discover and understand ghost-free higher dimensional generalizations of the NMG models, with no success. NMG corresponds to $\mathcal{L}_{(2)}^{\text{HD}}$ for $d = 3$ and $m_g c_R < 0$ in which case the theory contains a healthy massive spin-2 mode. With $m_g c_R < 0$ the massless spin-2 mode becomes a ghost, but in $d = 3$ this mode does not propagate. This way one ends up with a healthy theory of a massive spin-2 field. However, this mechanism of removing the ghost is peculiar to $d = 3$ and attempts to generalize NMG beyond $d = 3$ has so far remained largely unsuccessful [8]. The present analysis indicates that the bimetric theory is indeed one (possibly the only) minimal completion of NMG, providing a ghost-free generalization to all dimensions.

4 Conformal gravity from partially massless bimetric theory

Recently, the problem of finding nonlinear realizations of partial masslessness, first observed in the linear Fierz-Pauli theory in dS backgrounds [44], has attracted some attention [3, 4, 46, 49, 50, 55]. One such nonlinear candidate is the PM bimetric theory that exists only in 3 and 4 dimensions, and that directly contains the original linear PM theory in dS backgrounds [3, 4]. In these works it was not established that away from dS backgrounds this theory still contains six modes (in $d = 4$), rather than the generic seven modes of the HR bimetric theory. Another possible PM candidate is conformal gravity whose spectrum around dS backgrounds was observed to be similar to the linear PM spectrum [14]. CG in general propagates six modes, but the spectrum is plagued by the usual ghosts of HD gravity. On the face of it, these are very different approaches to the PM issue.

It is now easy to check that the HD gravity corresponding to the PM bimetric theory is indeed related to conformal gravity. What is more, it turns out that at the 4-derivative level, the PM bimetric theory and CG have exactly the same equations of motion (which is not the case for the non-PM actions considered above). Hence the PM bimetric theory is indeed a ghost-free extension of conformal gravity. Conversely, this provides further evidence that the PM bimetric theory propagates six modes even away from dS backgrounds.

4.1 The correspondence in $d = 4$

The partially massless bimetric theory in $d = 4$ is specified by setting [3, 4],

$$\beta_1 = \beta_3 = 0, \quad \alpha^4 \beta_0 = 3\alpha^2 \beta_2 = \beta_4. \quad (4.1)$$

Then, (2.18), (2.19) and (2.26) give,

$$a^2 = -\frac{1}{\alpha^2}, \quad s_1 = -\frac{2\beta_2}{\alpha^2}, \quad \alpha_0 = \alpha_1 = 0, \quad \alpha_2 = \frac{2\beta_2}{\alpha^2}. \quad (4.2)$$

and the solution for $f_{\mu\nu}$ takes the form,

$$f_{\mu\nu} = -\frac{1}{\alpha^2}g_{\mu\nu} + \frac{1}{\beta_2 m^2}P_{\mu\nu} + \mathcal{O}(m^{-4}). \quad (4.3)$$

Using these parameter values in (2.39), the higher derivative Lagrangian (2.38) reduces to,

$$\mathcal{L}^{\text{CG}} = -\frac{\alpha^2 m_g^2}{2\beta_2 m^2} \sqrt{g} \left(R^{\mu\nu} R_{\mu\nu} - \frac{1}{3} R^2 \right) + \mathcal{O}(m^{-4}). \quad (4.4)$$

Restricted to quadratic order in curvatures, this is the well-known action for conformal gravity. It is invariant under Weyl scalings of the metric and as a result it contains only 4 propagating massive modes, in addition to 2 massless modes.

The relation between conformal gravity and PM bimetric theory is constrained enough that the above reasoning can be carried out in reverse: demanding that the higher derivative action (2.38) reduces to conformal gravity at the quadratic level (that is, $\alpha_0 = 0$, $c_1 = 0$) fixes the parameters of the bimetric theory to their PM values. Explicitly, $\alpha_0 = 0$ and $s_0 = 0$ (2.19) imply $\alpha_1 = 0$, since $\alpha_0 - \alpha_1 = s_0$ (2.41). Hence, in order for the Einstein-Hilbert term to vanish, that is, $c_1 = 0$, we must have $a^2 = -\alpha^{-2}$. Then the $s_0 = 0$ equation (2.19) becomes

$$\beta_0 + \frac{3i}{\alpha}\beta_1 - \frac{3}{\alpha^2}\beta_2 - \frac{i}{\alpha^3}\beta_3 = 0. \quad (4.5)$$

Furthermore the condition $\alpha_0 = 0$ reads (2.26),

$$\beta_0 + \frac{4i}{\alpha}\beta_1 - \frac{6}{\alpha^2}\beta_2 - \frac{4i}{\alpha^3}\beta_3 + \frac{1}{\alpha^4}\beta_4 = 0. \quad (4.6)$$

The real and imaginary parts must vanish simultaneously which gives (4.1) as the unique solution. This establishes the correspondence between CG and PM bimetric theory.

4.2 A step further: Equivalence between CG and PM bimetric theory

As we have seen, in general, the bimetric equations of motion and the corresponding HD gravity equations differ by an operator $\delta f / \delta g$. On truncating to 4-derivative terms, as in $S_{(2)}^{\text{HD}}$, one finds that the truncated theories are not equivalent. An exception is the PM case as shown below.

The equations of motion arising from the conformal gravity action (1.3) correspond to setting the Bach tensor $B_{\mu\nu}$ of $g_{\mu\nu}$ to zero [9],

$$B_{\mu\nu} \equiv -\nabla^2 P_{\mu\nu} + \nabla^\rho \nabla_{(\mu} P_{\nu)\rho} + W_{\rho\mu\nu\sigma} P^{\rho\sigma} = 0. \quad (4.7)$$

In this expression the four-dimensional Weyl tensor $W_{\rho\mu\nu\sigma}$ is given by

$$W_{\rho\mu\nu\sigma} \equiv R_{\rho\mu\nu\sigma} + g_{\mu(\nu} R_{\sigma)\rho} - g_{\rho(\nu} R_{\sigma)\mu} + \frac{1}{3} g_{\rho(\nu} g_{\sigma)\mu} R. \quad (4.8)$$

Let us now consider the corresponding equation in the PM bimetric theory. The bimetric g -equation has already been solved to determine S (2.22) and $f_{\mu\nu}$ (2.31). We substitute these

in the bimetric f -equation (2.10) using the expressions (2.34) and (2.35) for the curvatures of $f_{\mu\nu}$. This gives a higher derivative equation for $g_{\mu\nu}$. On restricting to the PM parameter values (4.1), it is straightforward to show that the cosmological constant term and the Einstein tensor drop out and the equation reduces to,

$$-B_{\mu\nu} + \mathcal{O}(m^{-4}) = 0. \quad (4.9)$$

Hence the PM bimetric equation, evaluated to lowest non-trivial order in $R_{\mu\nu}(g)/m^2$, coincides with the CG equation.⁸

In other words, the conformal gravity equation (4.7) is the genuine low curvature limit of the PM bimetric theory. The difference is that while some of the propagating modes arising from this equation have negative norms in the CG framework, they appear as healthy non-ghost modes in the HR bimetric setup. This implies that the PM bimetric theory can be regarded as a ghost-free extension of conformal gravity.

This equivalence also has implications for the PM bimetric theory that has been identified as the only possible nonlinear generalization of the linear PM theory in the HR bimetric setup [3]. It is known that this theory has six propagating modes around dS backgrounds, a massless spin-2 field and the four-component PM field. The seventh mode, which exists in a generic HR bimetric theory, is eliminated by a new gauge invariance [44]. The question is if away from dS backgrounds this theory will still have a gauge invariance and, consequently, propagate only six modes. Some indirect evidence for this has been discussed in [3, 4]. The relation to conformal gravity found here provides direct evidence for the proposal as it establishes that, in the small curvature limit, the PM bimetric theory indeed propagates only six modes *around any background*.

The Bach equation (4.7) propagates six modes not only around dS [14] but also around flat background [13]. Thus an obvious question is whether the higher derivative corrections to the Bach equation in (4.9) can alter the propagator for the linear fluctuation around flat background. We will argue now that this is not the case and the propagator for the fluctuation $\delta g_{\mu\nu}$ around flat space is not affected by the higher derivative corrections to (4.9). Namely, consider the solution for $f_{\mu\nu}$ in terms of $g_{\mu\nu}$ given in (4.3). Taking into account the fact that the terms nonlinear in $P_{\mu\nu}(g)$ do not contribute to the linearized equation because $P_{\mu\nu}(\eta) = 0$, only two derivatives are acting on $\delta g_{\mu\nu}$. It is then also clear that when this solution is plugged into the $f_{\mu\nu}$ equation of motion, the resulting linearized equation will contain at most two more derivatives acting on $\delta g_{\mu\nu}$ coming from the Einstein tensor for $f_{\mu\nu}$. Hence the linearized equation for $g_{\mu\nu}$ around flat space is fourth order in derivatives. The fluctuation equation obtained in this way is of course the same as the linearized version of (4.9), in which all four-derivative terms come from the Bach tensor. Hence, the propagator around flat space is obtained from the Bach tensor alone and is therefore identical to the propagator in conformal

⁸This is of course consistent with the general analysis of the previous sections and follows from the fact that, in the PM case, only the lowest order term in $\delta f/\delta g \sim a^2 = -1/\alpha^2$ contributes to the relations between the equations of motion.

gravity. A similar argument is valid for dS backgrounds, confirming the result obtained in the bimetric formulation that a linear gauge symmetry is present around these backgrounds.

Also, in the small curvature limit, the symmetry transformations of the bimetric fields $g_{\mu\nu}$ and $f_{\mu\nu}$ follow from the Weyl invariance of conformal gravity. Thus to linear order in curvatures and in the gauge parameter, and modulo coordinate transformations, we get,

$$g'_{\mu\nu} = g_{\mu\nu} + \phi g_{\mu\nu}, \quad f'_{\mu\nu} = f_{\mu\nu} - \frac{\phi}{\alpha^2} g_{\mu\nu} - \frac{1}{\beta_2 m^2} \nabla_\mu^g \nabla_\nu^g \phi. \quad (4.10)$$

Let us now consider the relation between dS backgrounds in conformal gravity and in the PM bimetric theory. The CG equation (4.7) admits dS solutions for which $R_{\mu\nu}(\bar{g}) = \bar{g}_{\mu\nu} \Lambda$ (or $P_{\mu\nu}(\bar{g}) = \frac{1}{3} \bar{g}_{\mu\nu} \Lambda$). The cosmological constant is not determined by the theory and can be changed to any value by constant Weyl scalings of the metric, which are symmetries of the theory.

This solution of CG is mapped to bimetric theory by (4.3). It turns out that on dS backgrounds all higher curvature corrections to (4.3) vanish and the terms explicitly written out give,

$$\bar{f}_{\mu\nu} = \left(-\frac{1}{\alpha^2} + \frac{\Lambda}{3\beta_2 m^2} \right) \bar{g}_{\mu\nu}. \quad (4.11)$$

Writing this as $\bar{f}_{\mu\nu} = c^2 \bar{g}_{\mu\nu}$, we obtain,

$$\Lambda = m^2 \left(\frac{1 + c^2 \alpha^2}{\alpha^2} \right) 3\beta_2, \quad (4.12)$$

which is indeed the expression for the cosmological constant in terms of parameters of the PM bimetric theory. In this setup c is an arbitrary gauge dependent parameter. The flat space limit corresponds to $c^2 = -1/\alpha^2$. In the bimetric case, it is possible to express the theory in terms of variables such that all observables become independent of c [4].

Let's consider conformal gravity solutions away from dS spacetimes, that is, when $R_{\mu\nu}(g)$ is not proportional to $g_{\mu\nu}$. Such solutions, which are Bach flat but not conformally Einstein, are known to exist, see e.g. [59]. Then from (4.3) it follows that in the bimetric setup, the field $f_{\mu\nu}$ is not proportional to $g_{\mu\nu}$. In such cases the bimetric spectrum does not generically decompose into a massless and a massive spin-2 field [19]. Consequently, on the conformal gravity side too, the spectrum is not expected to follow such a decomposition away from dS backgrounds.

5 Discussions

Our results are summarized in the introduction section. Here we would like to first reemphasize two issues.

1. The analysis in this paper shows that the HR bimetric theory captures the essential features of higher derivative gravity actions (1.1) while at the same time avoiding the spin-2 ghost problem. The correspondence between the two theories found here is not

a complete equivalence of equations of motion, but can still be used to generate higher derivative completions of the 4-derivative gravity actions.

2. The equation of motion in the candidate PM bimetric theory at the 4-derivative level was shown to coincide with the Bach equation of conformal gravity. While this result was motivated by the general correspondence between bimetric and HD gravity actions, it turns out to bypass the general correspondence and, in fact, was an equivalence at the level of equations of motion. As a result it has genuine consequences for the bimetric PM proposal as discussed in the paper.

We also reiterate that the analysis of [40, 41], on acausal propagation and absence of a nonlinear PM symmetry in massive gravity, relies on the presence of a non-dynamical metric and, hence, breaks down for the HR bimetric theory. For this reason we have emphasized the distinctions between the HR bimetric models and the dRGT massive gravity models.

One may also consider alternative formulations and generalizations of the PM bimetric model in terms of vielbeins. Introducing vielbeins e^a_μ and \tilde{e}^a_μ for $g_{\mu\nu}$ and $f_{\mu\nu}$ respectively, the square-root matrix can be expressed as $(\sqrt{g^{-1}f})^\mu_\nu = e^\mu_c \tilde{e}^c_\nu$, provided the vielbeins satisfy the symmetrization condition, $\eta_{ac} \tilde{e}^c_\rho e^\rho_b = e_a{}^\rho \tilde{e}^c_\rho \eta_{cb}$, or in matrix form,

$$\eta \tilde{e} e^{-1} = (\eta \tilde{e} e^{-1})^T. \quad (5.1)$$

In general this condition can be implemented through a Lorentz transformation, though there are caveats [60] to this. However, it has also been shown that the bimetric theory expressed in terms of unconstrained vielbeins contains the above symmetrization condition as part of its equations of motion [61]. This work also formulates more general multivielbein theories that, with their equivalent metric formulations [62], open the way for investigating PM theories with more than one extra gauge invariance.

It would also be interesting to consider the conformal gravity analogues of the PM bimetric theory in higher dimensions. It is known that on adding Lanczos-Lovelock terms to the bimetric action, one can obtain candidate PM bimetric theories in $d = 6$ and $d = 8$ [4]. One then expects that on eliminating $f_{\mu\nu}$, the corresponding equations of motion, to lowest order in the curvature expansion, coincide with the equations of the recently constructed conformal gravity theories in 6 [63, 64] and 8 dimensions [65].

Acknowledgements:

We would like to thank Shailesh Lal and Bo Sundborg for discussions.

A Higher derivative treatment of free massive spin-0 and spin-2 fields

To better understand the relation between bimetric theory and higher derivative gravity and to resolve the ghost problem of the latter, we first illustrate the procedure in a much simpler scalar field analogue of the bimetric action. Then we consider the linearized bimetric theory to emphasize the parallel with the scalar field case.

A.1 Higher derivative treatment of massive scalars

Consider the theory of two scalar fields ϕ and ψ with a $\phi\psi$ mixing,

$$S = \int d^d x \left[-\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} \partial_\mu \psi \partial^\mu \psi - \frac{\mu^2}{2} (\phi + \psi)^2 - V(\phi, \psi) - \phi J_\phi - \psi J_\psi \right]. \quad (\text{A.1})$$

It describes a pair of massless and massive scalars, $\Phi_0 = \phi - \psi$, $\Phi_m = \phi + \psi$, in a non-diagonal basis. Conventionally this theory is analyzed in terms of the mass eigenstates. Here we want to discuss two approaches to recast (A.1) as a higher derivative theory of a single field, with emphasis on the ghost problem. But, for later reference, we first review the usual discussion of ghosts in terms of the vacuum persistence amplitude. For $V = 0$, the field equations are,

$$\frac{\delta S}{\delta \phi} = 0 \quad \Rightarrow \quad \square \phi - \mu^2 (\phi + \psi) = J_\phi, \quad (\text{A.2})$$

$$\frac{\delta S}{\delta \psi} = 0 \quad \Rightarrow \quad \square \psi - \mu^2 (\phi + \psi) = J_\psi, \quad (\text{A.3})$$

On adding and subtracting they diagonalize to,

$$\square \Phi_0 = J_0, \quad (\square - 2\mu^2) \Phi_m = J_m, \quad (\text{A.4})$$

with

$$\Phi_0 = \phi - \psi, \quad \Phi_m = \phi + \psi, \quad J_0 = J_\phi - J_\psi, \quad J_m = J_\phi + J_\psi. \quad (\text{A.5})$$

Hence, (A.1) describes a massless scalar Φ_0 and a massive one Φ_m of mass $2\mu^2$. In particular, the solution for $\phi = \frac{1}{2}(\Phi_0 + \Phi_m)$ is,⁹

$$\phi = \frac{1}{2} \left[\square^{-1} J_0 + (\square - 2\mu^2)^{-1} J_m + \Phi_0^{\text{hom}} + \Phi_m^{\text{hom}} \right], \quad (\text{A.6})$$

where $\Phi_{0,m}^{\text{hom}}$ are the homogeneous solutions of (A.4). From (A.1), on completing squares, one reads off the vacuum persistence amplitude,

$$A = \int \mathcal{D}\phi e^{iS} = \exp \left(-\frac{i}{4} \int d^d x \left[J_0(x) \square^{-1} J_0(x) + J_m(x) (\square - 2\mu^2)^{-1} J_m(x) \right] \right) < 1. \quad (\text{A.7})$$

The propagators $i\square^{-1}$ and $i(\square - 2\mu^2)^{-1}$ have positive residues at the corresponding mass poles which results in $A < 1$ meaning that the theory is ghost-free. In other words, the vacuum has decayed only into positive norm states due to interaction with the sources. In a theory that contains ghosts it becomes possible to get $A > 1$ implying a breakdown of unitarity and signalling vacuum decay into negative norm states. In the following we will use this criterion to check for ghosts in free higher derivative theories where A can be computed.

The potential $V(\phi, \psi)$ provides the interactions between the two mass eigenstates. At low energies, the massive mode Φ_m can be integrated out. In particular, for $V = 0$ this gives a Lagrangian for the free massless field Φ_0 with no knowledge of the massive mode. The analysis below does not involve integrating out fields in this way. Let us now consider the two approaches to obtaining higher derivative equations from (A.1).

⁹We use the standard notation, $(\square - m^2)^{-1} J(x) = \int d^d y G(x-y) J(y)$ with the Green's function $G(x-y) = (\square - m^2)^{-1} \delta(x-y) = -\int d^d k \frac{e^{ik(x-y)}}{k^2 + m^2 + i\epsilon}$.

First approach: the equivalent higher derivative equations:

Instead of diagonalizing the equations of motion (A.2) and (A.3), it is instructive to rewrite them as an equivalent higher derivative equation. The purpose is to illustrate that, given a higher derivative equation, the appearance of ghosts is not a forgone conclusion but depends on the structure of the source terms and on the choice of action.

Let us eliminate ψ between the equations of motion. The ϕ -equation (A.2) is algebraic in ψ and solves to,

$$\psi = -\phi + \frac{1}{\mu^2} (\Box\phi - J_\phi) . \quad (\text{A.8})$$

Such algebraic solutions can be obtained even for a non-zero $V(\phi, \psi)$, at least perturbatively in $1/\mu^2$. Inserting this in (A.3) gives a higher derivative equation,¹⁰

$$\Box (\Box - 2\mu^2) \phi = \frac{1}{2} [(\Box - 2\mu^2)J_0 + \Box J_m] . \quad (\text{A.9})$$

The specific source structure in (A.9), and the choice of action leading to it, are crucial for the absence of ghost. Had we ignored the sources in the above derivation and instead introduced a generic source coupling at the end, we would obtain $\Box (\Box - 2\mu^2) \phi = j$, corresponding to an action $\int d^d x [\frac{1}{2} \phi \Box (\Box - 2\mu^2) \phi - j\phi]$. This gives the correct mass poles, but implies that the massive mode is a ghost.¹¹ Of course, it is not obvious why two propagating modes of different masses should couple to the same source with the same strength.

However, with the source structure of (A.9) one gets the correct solution (A.6). To check for the presence of ghosts one needs to specify an action. The obvious procedure is to integrate out ψ in (A.1) (by completing the squares for ψ , or equivalently, eliminating ψ through the solution of *its own* equation of motion, and not (A.8)). Then for $V = 0$ one obtains the bi-local higher derivative action for ϕ ,

$$S_{bl} = \frac{1}{2} \int d^d x d^d y \left[\phi(x) G(x-y) \left[\Box_y (\Box_y - 2\mu^2) \phi(y) - (\Box_y - 2\mu^2) J_0(y) + \Box_y J_m(y) \right] \right. \\ \left. - \frac{1}{4} [J_m(x) - J_0(x)] G(x-y) [J_m(y) - J_0(y)] \right] . \quad (\text{A.10})$$

By construction, this has (A.9) as equation of motion and gives the vacuum persistence amplitude (A.7), consistent with the absence of ghosts. Although the $J - J$ term in the second line does not contribute to the equation of motion, it is needed to get the correct ghost-free amplitude. Its pretense simply indicates that the interactions of the ψ field cannot all be encoded in the structure of the ϕ action alone. To summarize, the above illustrates that a higher-derivative equation like $\Box(\Box - 2\mu^2)\phi = 0$, which normally implies a ghost, could actually be ghost-free with the appropriate action and source couplings.

¹⁰The same result follows on first solving the (A.3) for ψ and substituting the outcome in (A.2).

¹¹In the naive propagator for this theory, $\frac{1}{k^2(k^2+m^2)} = \frac{1}{m^2}(\frac{1}{k^2} - \frac{1}{k^2+m^2})$, the residues of the massless and massive poles have different signs so in the analogue of (A.7), A is not always bounded by unity.

Second approach: a more general higher derivative action:

Now, instead of using the algebraic solution $\psi(\phi, J_\phi)$ to eliminate ψ from the equations of motion (first approach), we use it in the action to get $S[\phi, \psi(\phi, J_\phi)]$. This gives a 6-derivative equation of motion for ϕ which is more general and contains the solutions of (A.9) as a subsector. The analogue of this treatment, when applied to bimetric theory, reproduces the higher derivative gravity theories. The discussion here is intended to shed light on the emergence of the spin-2 ghost in higher derivative gravity.

Using the solution (A.8) to eliminate ψ in the action (A.1) gives a 6-derivative action,

$$S[\phi, \psi(\phi, J_\phi)] = \frac{1}{2\mu^4} \int d^d x \left[\phi \square (\square - \mu^2) (\square - 2\mu^2) \phi - 2\phi (\square - \mu^2) ((\square - \mu^2) J_\phi + \mu^2 J_\psi) \right. \\ \left. + J_\phi (\square - \mu^2) J_\phi + 2\mu^2 J_\phi J_\psi \right]. \quad (\text{A.11})$$

The corresponding equation of motion is,

$$\left. \frac{\delta S}{\delta \phi(x)} \right|_\psi + \int d^d y \frac{\delta \psi(y)}{\delta \phi(x)} \left. \frac{\delta S}{\delta \psi(y)} \right|_\phi = 0. \quad (\text{A.12})$$

The first term vanishes by virtue of (A.8). Hence one gets, $\int d^d y \frac{\delta \psi(y)}{\delta \phi(x)} \frac{\delta S}{\delta \psi(y)} \Big|_\phi = 0$ which differs from $\frac{\delta S}{\delta \psi} = 0$ by the extra operator, $\frac{\delta \psi(y)}{\delta \phi(x)} = \frac{1}{\mu^2} (\square_y - \mu^2) \delta(x - y)$ as follows from (A.8). This gives the higher derivative equation,

$$(\square - \mu^2) \left[\square (\square - 2\mu^2) \phi - \frac{1}{2} ((\square - 2\mu^2) J_0 + \square J_m) \right] = 0. \quad (\text{A.13})$$

The terms within square brackets are the same as in (A.9). Had we ignored the sources from the beginning and then inserted a generic source coupling at the end we would obtain $\square(\square - \mu^2)(\square - 2\mu^2)\phi = j$ corresponding to an action $\int d^4 x \frac{1}{2} \phi \square (\square - \mu^2) (\square - 2\mu^2) \phi - j\phi$. This would give a theory with three mass poles and a ghost mode. However, for the correct source structure, (A.13) is equivalent to

$$\square (\square - 2\mu^2) \phi - \frac{1}{2} ((\square - 2\mu^2) J_0 + \square J_m) = \chi^{\text{hom}}, \quad (\text{A.14})$$

where χ^{hom} solves $(\square - \mu^2)\chi^{\text{hom}} = 0$ and needs to be specified through two boundary conditions. The boundary conditions that give $\chi^{\text{hom}} = 0$ specify the subsector of (A.11) that is equivalent to (A.1).

In general, the complete solution becomes,

$$\phi = \frac{1}{2} \left[\square^{-1} J_0 + (\square - 2\mu^2)^{-1} J_m \right] + \square^{-1} (\square - 2\mu^2)^{-1} \chi^{\text{hom}} + \frac{1}{2} \left[\Phi_0^{\text{hom}} + \Phi_m^{\text{hom}} \right]. \quad (\text{A.15})$$

A non-vanishing χ^{hom} appears as an arbitrariness in the sources J_0 and J_m , but even then only the two original mass poles of (A.1) contribute to the solution.

To address the ghost issue one can again compute the vacuum persistence amplitude $\int \mathcal{D}\phi e^{iS}$ by completing the squares for ϕ in the action (A.11). The answer turns out to be exactly the same as (A.7). In particular, there is no contribution from χ^{hom} . Hence the complete untruncated higher derivative theory (A.11) does not propagate ghosts.

Truncation to a 4 derivative theory:

In the action (A.11) the highest derivative terms come with $1/\mu^4$. Dropping these results in a 4-derivative action for ϕ with mass poles at $m^2 = 0$ and $m^2 = 2\mu^2/3$. On completing the squares, one can see that the massless mode is healthy while the massive mode is now a ghost. Thus the truncation has shifted the massive pole away from $2\mu^2$ and turned it into a ghost. Of course, in this example we know that the theory cannot be trusted near the massive pole which is close to the truncation scale. Near the massive pole, higher derivative corrections become important and including these both shifts the mass pole to its correct value as well as eliminates the ghost.

The analogue of (A.11) in the bimetric setup is given as an expansion in powers of derivatives, suppressed by the mass parameter. When truncated to 4-derivative terms, the massive spin-2 field becomes a ghost with a shifted mass, as an artifact of the truncation. The lesson is that the spin-2 ghost in HD gravity must be understood in a similar way, although in the presence of interactions, the complete HD theory will not have finite number of terms.

A.2 Higher derivative treatment of linearized bimetric theory

Let us now repeat this procedure for spin-2 fluctuations on a curved background $\bar{g}_{\mu\nu}$. The notation here simplifies and streamlines the manipulations. Indices are raised using \bar{g} . Our starting point is the quadratic Fierz-Pauli action with linear interaction given by,

$$S = \int d^d x \sqrt{\bar{g}} \left[\frac{1}{2} h_{\mu\nu} D^{\mu\nu\rho\sigma}(\mu^2, \Lambda) h_{\rho\sigma} + \frac{1}{2} \tilde{h}_{\mu\nu} D^{\mu\nu\rho\sigma}(\tilde{\mu}^2, \tilde{\Lambda}) \tilde{h}_{\rho\sigma} + \lambda h_{\mu\nu} G_{(1)}^{\mu\nu\rho\sigma} \tilde{h}_{\rho\sigma} - h_{\mu\nu} T^{\mu\nu} - \tilde{h}_{\mu\nu} \tilde{T}^{\mu\nu} \right], \quad (\text{A.16})$$

where we have defined

$$D^{\mu\nu\rho\sigma}(\mu^2, \Lambda) = \mathcal{E}^{\mu\nu\rho\sigma} + \frac{\mu^2}{2} G_{(1)}^{\mu\nu\rho\sigma} - \frac{\Lambda}{d-2} G_{(1/2)}^{\mu\nu\rho\sigma}, \quad (\text{A.17})$$

in terms of the linear Einstein operator $\mathcal{E}^{\mu\nu\rho\sigma}$ and the generalized DeWitt metrics, which we define through

$$G_{(\xi)}^{\mu\nu\rho\sigma} = \bar{g}^{\mu\nu} \bar{g}^{\rho\sigma} - \xi \bar{g}^{\mu(\rho} \bar{g}^{\sigma)\nu}. \quad (\text{A.18})$$

The field equations are given by

$$D^{\mu\nu\rho\sigma}(\mu^2, \Lambda) h_{\rho\sigma} + \lambda G_{(1)}^{\mu\nu\rho\sigma} \tilde{h}_{\rho\sigma} - T^{\mu\nu} = 0, \quad (\text{A.19})$$

$$D^{\mu\nu\rho\sigma}(\tilde{\mu}^2, \tilde{\Lambda}) \tilde{h}_{\rho\sigma} + \lambda G_{(1)}^{\mu\nu\rho\sigma} h_{\rho\sigma} - \tilde{T}^{\mu\nu} = 0, \quad (\text{A.20})$$

Restricting now to $\tilde{\Lambda} = \Lambda$ and $\tilde{\mu}^2 = 2\lambda = \mu^2$ we can again simply add and subtract these equations to get

$$D^{\mu\nu\rho\sigma}(2\mu^2, \Lambda) (h_{\rho\sigma} + \tilde{h}_{\rho\sigma}) - (T^{\mu\nu} + \tilde{T}^{\mu\nu}) = 0, \quad (\text{A.21})$$

$$D^{\mu\nu\rho\sigma}(0, \Lambda) (h_{\rho\sigma} - \tilde{h}_{\rho\sigma}) - (T^{\mu\nu} - \tilde{T}^{\mu\nu}) = 0. \quad (\text{A.22})$$

These are the equations of a massive spin-2 field $h_{\mu\nu}^+ = h_{\mu\nu} + \tilde{h}_{\mu\nu}$ with a mass $2\mu^2$ and a massless spin-2 field $h_{\mu\nu}^- = h_{\mu\nu} - \tilde{h}_{\mu\nu}$ on a curved background. We can solve for $\tilde{h}_{\mu\nu}$ in (A.19), to get

$$G_{(1)}^{\mu\nu\rho\sigma}\tilde{h}_{\rho\sigma} = -\frac{2}{\mu^2} [D^{\mu\nu\rho\sigma}(\mu^2, \Lambda)h_{\rho\sigma} - T^{\mu\nu}] , \quad (\text{A.23})$$

and subsequently

$$\tilde{h}_{\mu\nu} = \frac{2}{\mu^2} \left[D_{\mu\nu}{}^{\rho\sigma}(\mu^2, \Lambda)h_{\rho\sigma} - T_{\mu\nu} - \frac{1}{d-1}\bar{g}_{\mu\nu} (D^\alpha{}_\alpha{}^{\rho\sigma}(\mu^2, \Lambda)h_{\rho\sigma} - T) \right] . \quad (\text{A.24})$$

Inserting this solution in (A.20) and using some basic manipulations (see the end of this section) of the DeWitt metrics one obtains the higher derivative equation

$$D^{\mu\nu\alpha\beta}(0, \Lambda)G_{\alpha\beta\rho\sigma}^{(d-1)}D^{\rho\sigma\lambda\kappa}(2\mu^2, \Lambda)h_{\lambda\kappa} - \left(D^{\mu\nu\alpha\beta}(\mu^2, \Lambda)G_{\alpha\beta\rho\sigma}^{(d-1)}T^{\rho\sigma} - \frac{d-1}{2}\mu^2\tilde{T}^{\mu\nu} \right) = 0 . \quad (\text{A.25})$$

Again, the structure of the higher derivative operator hitting $h_{\lambda\kappa}$ is the composite of a zero mass operator and a massive operator for a mass $2\mu^2$. The contraction of these operators is of course a bit more complex now, involving a DeWitt metric. In fact this is the inverse of $G_{(1)}$, in the sense that

$$G_{(1)}^{\mu\nu\rho\sigma}G_{\alpha\beta\rho\sigma}^{(d-1)} = \frac{d-1}{2} \left(\delta_\alpha^\mu\delta_\beta^\nu + \delta_\beta^\mu\delta_\alpha^\nu \right) . \quad (\text{A.26})$$

In the source contribution of (A.25), one of the sources is contracted with an operator $\sim \delta\tilde{h}^{\mu\nu}/\delta h_{\rho\sigma}$. All of this is in complete analogy with the scalar field example and generalizes that discussion in a very straightforward way, the only complication being the metric of contraction. The insertion of the solutions (A.23) and (A.24) into the action also proceeds in an analogue fashion and again results in an operator $D(\mu^2, \Lambda)$ contracting (A.25) with the DeWitt metric $G^{(d-1)}$. This makes it obvious that the implications of absence of ghost for the untruncated theory go through just as in the scalar field example and we do not provide the details of this here.

In order to check the details at the level of the action it is useful to observe that

$$G_{(1)}^{\mu\nu\alpha\beta}G_{(1)\alpha\beta}{}^{\rho\sigma} + G_{(1)}^{\mu\nu\rho\sigma} - \frac{1}{d-1}G_{(1)}^{\mu\nu\alpha}G_{(1)\beta}{}^{\rho\sigma} = 0 , \quad (\text{A.27})$$

and

$$G_{(1)\alpha}{}^{\rho\sigma} = G_{(1)}^{\rho\sigma\alpha} = (d-1)\bar{g}^{\rho\sigma} , \quad (\text{A.28})$$

together with (for any D)

$$D^{\alpha\beta\rho\sigma}G_{(1)\alpha\beta}^{\mu\nu} = D^\alpha{}_\alpha{}^{\rho\sigma}\bar{g}^{\mu\nu} - D^{\mu\nu\rho\sigma} , \quad D^{\mu\nu}{}_{\alpha\beta}G_{(1)}^{\alpha\beta\rho\sigma} = D^{\mu\nu\alpha}{}_\alpha\bar{g}^{\rho\sigma} - D^{\mu\nu\rho\sigma} . \quad (\text{A.29})$$

These results can be used to show that

$$\begin{aligned} & D^{\mu\nu\alpha\beta}(\mu^2, \Lambda)D_{\alpha\beta}{}^{\rho\sigma}(\mu^2, \Lambda) - \frac{1}{d-1}D^{\mu\nu\alpha}{}_\alpha(\mu^2, \Lambda)D^\beta{}_\beta{}^{\rho\sigma}(\mu^2, \Lambda) + \frac{\mu^4}{4}G_{(1)}^{\mu\nu\rho\sigma} \\ & = D^{\mu\nu\alpha\beta}(0, \Lambda)D_{\alpha\beta}{}^{\rho\sigma}(2\mu^2, \Lambda) - \frac{1}{d-1}D^{\mu\nu\alpha}{}_\alpha(0, \Lambda)D^\beta{}_\beta{}^{\rho\sigma}(2\mu^2, \Lambda) . \end{aligned} \quad (\text{A.30})$$

Together with the solutions (A.23) and (A.24), these relations make the manipulations straightforward.

B The general perturbative solution of the $g_{\mu\nu}$ equation for $f_{\mu\nu}$

For general β_n finding an exact solution for $S = \sqrt{g^{-1}f}$ is not feasible.¹² But one can find a perturbative solution in powers of $\frac{1}{m^2}$. The equation of motion for $g_{\mu\nu}$ in terms of its Schouten tensor $P_{\mu\nu}$ becomes

$$\text{Tr} P \delta_\nu^\mu - P_\nu^\mu = m^2 \sum_{n=0}^{d-1} (-1)^n \beta_n \mathbb{Y}_{(n)\nu}^\mu(S), \quad (\text{B.1})$$

where the matrices $\mathbb{Y}_{(n)}(S)$ are defined in (2.11), and here and in the following we raise and lower indices with $g_{\mu\nu}$. Our goal is to solve the equations (B.1) for S_ν^μ perturbatively in P_ν^μ/m^2 . Since the solution for S_ν^μ is a polynomial in P_ν^μ , the resulting action for $g_{\mu\nu}$ will be a higher curvature theory with coefficients given in terms of the parameters of the bimetric model. In order to find the perturbative solution for S_ν^μ , consider the general ansatz

$$\begin{aligned} S_\nu^\mu &= a \delta_\nu^\mu + \frac{1}{m^2} (b_1 P_\nu^\mu + b_2 \text{Tr} P \delta_\nu^\mu) \\ &+ \frac{1}{m^4} (c_1 P_\nu^{\mu 2} + c_2 P_\nu^\mu \text{Tr} P + c_3 \text{Tr}(P^2) \delta_\nu^\mu + c_4 (\text{Tr} P)^2 \delta_\nu^\mu) + \mathcal{O}(m^{-6}). \end{aligned} \quad (\text{B.2})$$

When we plug this into (B.1), the coefficients of different terms δ_ν^μ , P_ν^μ , $\delta_\nu^\mu \text{Tr} P$, $P_\nu^{\mu 2}$, \dots have to vanish separately. This will determine all the coefficients a, b_i, c_i in the ansatz for S_ν^μ .

To simplify the calculation, define $M_\nu^\mu \equiv S_\nu^\mu/a - \delta_\nu^\mu$ or, $S_\nu^\mu = a(\delta_\nu^\mu + M_\nu^\mu)$. Then, using

$$e_k(S) = a^k \sum_{m=0}^k \binom{d-m}{k-m} e_m(M), \quad (S^{n-k})_\nu^\mu = a^{n-k} \sum_{r=0}^{n-k} \binom{n-k}{r} (M^r)^\mu_\nu, \quad (\text{B.3})$$

the equations of motion can be written as

$$\text{Tr} P \delta_\nu^\mu - P_\nu^\mu = m^2 \sum_{n=0}^{d-1} \beta_n a^n \sum_{k=0}^n \sum_{m=0}^k \sum_{r=0}^{n-k} (-1)^{n+k} \binom{d-m}{k-m} \binom{n-k}{r} e_m(M) (M^r)^\mu_\nu. \quad (\text{B.4})$$

Since M_ν^μ starts with m^{-2} , a power of M^l will contribute to orders m^{-2l} and higher. In the following we work out the first three orders explicitly. To simplify the notation, we use the sums s_k defined in (2.18) that will be frequently encountered below.

Zeroth order: The equation of motion at $\mathcal{O}(m^0)$ is obtained from terms with the last two sums in (B.4) restricted to $(r, m) = (0, 0)$. This gives $s_0 \equiv \sum_{n=0}^{d-1} \binom{d-1}{n} \beta_n a^n = 0$ (2.19). This is a polynomial equation in a that determines it in terms of the β_n .

¹²In principle, exact solutions can be obtained implicitly by considering higher powers of the $g_{\mu\nu}$ -equation when written in the form $R_\nu^\mu = \sum_r q_r (S^r)^\mu_\nu$, where q_r are functions of β_n and $e_n(S)$. On tracing and expressing $\text{Tr}(S^r)$ for $r > d$ in terms $\text{Tr}(S^n)$ with $n \leq d$, one can obtain enough equations to determine all $e_n(S)$. Then the g -equation can be used to determine S in terms of the curvatures.

First order: At $\mathcal{O}(m^{-2})$ terms with $(r, m) = (1, 0)$ and $(r, m) = (0, 1)$ contribute in (B.4). Using simple identities for the binomial coefficients we obtain for this order

$$\text{Tr}P \delta_\nu^\mu - P_\nu^\mu = \frac{1}{a} \left[((d-1)b_2 + b_1) \text{Tr}P \delta_\nu^\mu - b_1 P_\nu^\mu \right] s_1, \quad (\text{B.5})$$

where s_1 is defined in (2.18). Comparing the coefficients of $\text{Tr}P \delta_\nu^\mu$ and P_ν^μ separately, we find the following equations that determines b_1 and b_2 in terms of the β_n .

$$b_2 = 0, \quad b_1 = \frac{a}{s_1}. \quad (\text{B.6})$$

Second order: At order $\mathcal{O}(m^{-4})$ we can have contributions from the following combinations of (r, m) ,

$$(1, 0), \quad (2, 0), \quad (1, 1), \quad (0, 1), \quad (0, 2). \quad (\text{B.7})$$

Demanding the coefficients of the different terms at this order to vanish separately then gives the following system of equations for the coefficients c_i in the ansatz (B.2),

$$\begin{aligned} s_1 c_1 - \frac{b_1^2 s_2}{a} &= 0, & s_1 c_2 + \frac{b_1^2 s_2}{a} &= 0, \\ s_1(c_1 + d c_3) - s_1 c_3 - \frac{b_1^2 s_2}{2a} &= 0, & s_1(c_2 + d c_4) - s_1 c_4 + \frac{b_1^2 s_2}{2a} &= 0. \end{aligned} \quad (\text{B.8})$$

The solution is easily found to be

$$c_1 = -c_2 = \frac{b_1^2 s_2}{a s_1} = \frac{a s_2}{s_1^3}, \quad c_3 = -c_4 = \frac{c_2}{2(d-1)} = -\frac{a s_2}{2(d-1) s_1^3}, \quad (\text{B.9})$$

where we have also used the solution for b_1 given in (2.20). Putting everything together, the solution for S_ν^μ to $\mathcal{O}(m^{-4})$, becomes (2.22). The procedure can straightforwardly be continued to compute the solution for S_ν^μ to any order in m^{-2} .

References

- [1] S. F. Hassan and R. A. Rosen, JHEP **1202** (2012) 126 [arXiv:1109.3515 [hep-th]].
- [2] S. F. Hassan and R. A. Rosen, JHEP **1204** (2012) 123 [arXiv:1111.2070 [hep-th]].
- [3] S. F. Hassan, A. Schmidt-May and M. von Strauss, arXiv:1208.1797 [hep-th].
- [4] S. F. Hassan, A. Schmidt-May and M. von Strauss, arXiv:1212.4525 [hep-th].
- [5] K. S. Stelle, Phys. Rev. D **16** (1977) 953.
- [6] K. S. Stelle, Gen. Rel. Grav. **9** (1978) 353.
- [7] E. A. Bergshoeff, O. Hohm and P. K. Townsend, Phys. Rev. Lett. **102** (2009) 201301 [arXiv:0901.1766 [hep-th]].
- [8] A. Kleinschmidt, T. Nutma, A. Virmani and , Gen. Rel. Grav. **45** (2013) 727 [arXiv:1206.7095 [hep-th]].

- [9] R. Bach, *Math. Zeitschr.* **9**, (1921) 110
- [10] M. Kaku, P. K. Townsend and P. van Nieuwenhuizen, *Phys. Lett. B* **69** (1977) 304.
- [11] E. S. Fradkin and A. A. Tseytlin, *Nucl. Phys. B* **201** (1982) 469.
- [12] S. C. Lee and P. van Nieuwenhuizen, *Phys. Rev. D* **26** (1982) 934.
- [13] R. J. Riegert, *Phys. Lett. A* **105** (1984) 110.
- [14] J. Maldacena, arXiv:1105.5632 [hep-th].
- [15] R. R. Metsaev, *JHEP* **1206** (2012) 062 [arXiv:0709.4392 [hep-th]].
- [16] P. D. Mannheim, *Found. Phys.* **42** (2012) 388 [arXiv:1101.2186 [hep-th]].
- [17] H. -J. Schmidt, *eConf C* **0602061** (2006) 12 [*Int. J. Geom. Meth. Mod. Phys.* **4** (2007) 209] [gr-qc/0602017].
- [18] S. Alexandrov, K. Krasnov, S. Speziale and , arXiv:1212.3614 [hep-th].
- [19] S. F. Hassan, A. Schmidt-May and M. von Strauss, arXiv:1208.1515 [hep-th].
- [20] M. S. Volkov, *JHEP* **1201** (2012) 035 [arXiv:1110.6153 [hep-th]].
- [21] M. von Strauss, A. Schmidt-May, J. Enander, E. Mortsell and S. F. Hassan, *JCAP* **1203** (2012) 042 [arXiv:1111.1655 [gr-qc]].
- [22] D. Comelli, M. Crisostomi, F. Nesti and L. Pilo, *JHEP* **1203** (2012) 067 [Erratum-ibid. **1206** (2012) 020] [arXiv:1111.1983 [hep-th]].
- [23] M. Berg, I. Buchberger, J. Enander, E. Mortsell and S. Sjors, arXiv:1206.3496 [gr-qc].
- [24] M. Park and L. Sorbo, arXiv:1212.2691 [hep-th].
- [25] Y. Sakakihara, J. Soda, T. Takahashi and , arXiv:1211.5976 [hep-th].
- [26] Y. Akrami, T. S. Koivisto, M. Sandstad and , *JHEP* **03** (2013) 099 [*JHEP* **1303** (2013) 099] [arXiv:1209.0457 [astro-ph.CO]].
- [27] S. Capozziello, P. Martin-Moruno and , *Phys. Lett. B* **719** (2013) 14 [arXiv:1211.0214 [gr-qc]].
- [28] M. Mohseni, *JCAP* **1211** (2012) 023 [arXiv:1211.3501 [hep-th]].
- [29] V. Baccetti, P. Martin-Moruno, M. Visser and , *JHEP* **1208** (2012) 108 [arXiv:1206.4720 [gr-qc]].
- [30] V. Baccetti, P. Martin-Moruno, M. Visser and , *JHEP* **1208** (2012) 148 [arXiv:1206.3814 [gr-qc]].
- [31] V. Baccetti, P. Martin-Moruno, M. Visser and , *Class. Quant. Grav.* **30** (2013) 015004 [arXiv:1205.2158 [gr-qc]].
- [32] M. S. Volkov, *Phys. Rev. D* **85** (2012) 124043 [arXiv:1202.6682 [hep-th]].
- [33] R. Myrzakulov, M. Shahalam and , arXiv:1303.0194 [gr-qc].
- [34] K. -i. Maeda, M. S. Volkov and , arXiv:1302.6198 [hep-th].
- [35] C. de Rham and G. Gabadadze, *Phys. Rev. D* **82** (2010) 044020 [arXiv:1007.0443 [hep-th]].
- [36] C. de Rham, G. Gabadadze and A. J. Tolley, *Phys. Rev. Lett.* **106** (2011) 231101 [arXiv:1011.1232 [hep-th]].

- [37] S. F. Hassan and R. A. Rosen, Phys. Rev. Lett. **108** (2012) 041101 [arXiv:1106.3344 [hep-th]].
- [38] S. F. Hassan, R. A. Rosen and A. Schmidt-May, JHEP **1202** (2012) 026 [arXiv:1109.3230 [hep-th]].
- [39] S. F. Hassan, A. Schmidt-May and M. von Strauss, *In preparation*
- [40] S. Deser and A. Waldron, Phys. Rev. Lett. 110, **111101** (2013) [arXiv:1212.5835 [hep-th]].
- [41] S. Deser, M. Sandora and A. Waldron, arXiv:1301.5621 [hep-th].
- [42] C. de Rham, K. Hinterbichler, R. A. Rosen and A. J. Tolley, arXiv:1302.0025 [hep-th].
- [43] A. Higuchi, Nucl. Phys. B **282** (1987) 397.
- [44] S. Deser and A. Waldron, Nucl. Phys. B **607** (2001) 577 [hep-th/0103198].
- [45] D. Francia, J. Mourad, A. Sagnotti and , Nucl. Phys. B **804** (2008) 383 [arXiv:0803.3832 [hep-th]].
- [46] E. Joung, L. Lopez and M. Taronna, JHEP **1207** (2012) 041 [arXiv:1203.6578 [hep-th]].
- [47] E. Joung, L. Lopez and M. Taronna, JHEP **1301** (2013) 168 [arXiv:1211.5912 [hep-th]].
- [48] Y. . M. Zinoviev, arXiv:1302.1983 [hep-th].
- [49] Y. .M. Zinoviev, Nucl. Phys. B **770** (2007) 83 [hep-th/0609170].
- [50] S. Deser, E. Joung and A. Waldron, Phys. Rev. D **86** (2012) 104004 [arXiv:1208.1307 [hep-th]].
- [51] D. G. Boulware and S. Deser, Phys. Rev. D **6** (1972) 3368.
- [52] A. De Felice, A. E. Gumrukcuoglu, C. Lin, S. Mukohyama and , arXiv:1303.4154 [hep-th].
- [53] S. F. Hassan, R. A. Rosen, JHEP **1107** (2011) 009. [arXiv:1103.6055 [hep-th]].
- [54] S. F. Hassan, A. Schmidt-May and M. von Strauss, Phys. Lett. B **715** (2012) 335 [arXiv:1203.5283 [hep-th]].
- [55] C. de Rham and S. Renaux-Petel, JCAP **1301** (2013) 035 [arXiv:1206.3482 [hep-th]].
- [56] M. F. Paulos, A. J. Tolley and , JHEP **1209** (2012) 002 [arXiv:1203.4268 [hep-th]].
- [57] M. S. Volkov, Phys. Rev. D **86** (2012) 104022 [arXiv:1207.3723 [hep-th]].
- [58] R. M. Wald, Chicago, Usa: Univ. Pr. (1984) 491p
- [59] H. -S. Liu, H. Lu, C. N. Pope, J. Vazquez-Poritz and , arXiv:1303.5781 [hep-th].
- [60] C. Deffayet, J. Mourad, G. Zahariade and , arXiv:1208.4493 [gr-qc].
- [61] K. Hinterbichler, R. A. Rosen and , JHEP **1207** (2012) 047 [arXiv:1203.5783 [hep-th]].
- [62] S. F. Hassan, A. Schmidt-May, M. von Strauss and , arXiv:1204.5202 [hep-th].
- [63] L. Bonora, P. Pasti, M. Bregola and , Class. Quant. Grav. **3** (1986) 635.
- [64] R. R. Metsaev, J. Phys. A **44** (2011) 175402 [arXiv:1012.2079 [hep-th]].
- [65] N. Boulanger, J. Erdmenger and , Class. Quant. Grav. **21** (2004) 4305 [hep-th/0405228].